Extreme suppression of lateral floret development by a single amino acid change in the VRS1 transcription factor

Sakuma, Shun; Lundqvist, Udda; Kakei, Yusuke; Thirulogachandar, Venkatasubbu; Suzuki, Takako; Hori, Kiyosumi

Published in:
Plant Physiology

DOI:
10.1104/pp.17.01149

Publication date:
2017

Document Version
Peer reviewed version

Plant Physiology Preview. Published on November 3, 2017, as DOI:10.1104/pp.17.01149

Running Title: *deficiens* barley

Corresponding author details:
Shun Sakuma, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, OT Gatersleben, D-06466 Stadt Seeland, Germany. Phone: +49 39482 5-519, Email: sakuma@ipk-gatersleben.de
Takao Komatsuda, Institute of Crop Science, National Agriculture and Food Research Organization (NARO), Kannondai 2-1-2, Tsukuba 305 8518 Japan. Phone: +81 29-838-7408, E-mail: takao@affrc.go.jp

Extreme suppression of lateral floret development by a single amino acid change in the VRS1 transcription factor

Shun Sakuma[1, 2, 3]*, Udda Lundqvist[4], Yusuke Kakei[5], Venkatasubbu Thirulogachandar[2], Takako Suzuki[6], Kiyosumi Hori[1, 7], Jianzhong Wu[1, 7], Akemi Tagiri[1], Twan Rutten[2], Ravi Koppolu[2], Yukihisa Shimada[5], Kelly Houston[8], William T. B. Thomas[8], Robbie Waugh[8, 9], Thorsten Schnurbusch[2], Takao Komatsuda[1, 7]*

[1]National Institute of Agrobiological Sciences (NIAS), Kannondai 2-1-2, Tsukuba 305 8602 Japan; [2]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, OT Gatersleben, D-06466 Stadt Seeland, Germany; [3]Faculty of Agriculture, Tottori University, Koyama-Minami 4-101, Tottori 680 8550, Japan; [4]Nordic Genetic Resource Center (NordGen), SE-23053 Alnarp, Sweden; [5]Yokohama City University, Maioka 640-12, Yokohama 244 0813, Japan; [6]Agricultural Research Department, Hokkaido Research Organization, Chuo Agricultural Experiment Station, Higashi 6, Kita 15, Naganuma, Hokkaido 069 1395, Japan; [7]Institute of Crop Science, National Agriculture and Food Research Organization (NARO), Kannondai 2-1-2, Tsukuba 305 8518 Japan; [8]The James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK, [9]Division of Plant Sciences, University of Dundee, Dundee DD1 4HN

***Corresponding authors:**

Shun Sakuma: sakuma@ipk-gatersleben.de

Takao Komatsuda: takao@affrc.go.jp

One sentence summary:

Extreme suppression of lateral floret development in *deficiens* barley is the result of a single amino acid substitution in the homeodomain-leucine zipper class I transcription factor VRS1.

Author contributions:

S.S. and T.K. designed research. S.S., U.L., and R.K. analyzed mutants. S.S., T. Suzuki, K.Hori, and J.W. carried out the fine mapping. S.S. and A.T. performed RNA in situ hybridization and qRT-PCR. S.S., Y.K., V.T., and Y.S. analyzed RNA-seq data. S.S. and T.R. performed SEM analysis. K.Houston analyzed the

European Winter barely accessions. S.S., W.T., R.W., T. Schnurbusch, and T.K. wrote the manuscript. All authors have reviewed and commented on the manuscript. The authors declare no conflict of interest.

Funding information:

This research was funded by the Ministry of Agriculture, Forestry, and Fisheries of Japan (Genomics for Agricultural Innovation grants no. TRS1002 to T.K. and S.S.); a Grant-in-Aid from the Japan Society for the Promotion of Science (JSPS) Postdoctoral Fellow for Research Abroad (to S.S.); a Grant-in-Aid for Young Scientists (B) (no. 16K18635 to S.S.). The field trials work was conducted under a work package of the EU FP7 ‘Improving nutrient use efficiency in major European food, feed and biofuel crops to reduce negative environmental impacts of crop production’ (NUE Crops) - EU-FP7 222-645 (2009 - 2014) led by Carlo Leifert of the University of Newcastle.

Abstract

Increasing grain yield is an endless challenge for cereal crop breeding. In barley, grain number is mainly controlled by *Six-rowed spike 1* (*Vrs1*) that encodes a homeodomain leucine zipper class I transcription factor. However, little is known about the genetic basis of grain size. Here we show that extreme suppression of lateral florets contributes to enlarged grains in *deficiens* barley. Through a combination of fine mapping and resequencing *deficiens* mutants we have identified that a single amino acid substitution at a putative phosphorylation site in VRS1 is responsible for the *deficiens* phenotype. *deficiens* mutant alleles confer an increase in grain size, reduction in plant height and a significant increase in thousand grain weight in contemporary cultivated germplasm. Haplotype analysis revealed that barley carrying the *deficiens* allele (*Vrs1.t1)* originated from two-rowed types carrying the *Vrs1.b2* allele, predominantly found in germplasm from Northern Africa. *In situ* hybridization of *histone H4*, a marker for cell cycle or proliferation, showed weaker expression in the lateral spikelets compared to central spikelets in *deficiens*. Transcriptome analysis revealed that a number of *histone* superfamily genes were upregulated in the *deficiens* mutant suggesting that enhanced cell proliferation in the central spikelet may contribute to larger grains. Our data suggest that grain yield can be improved by suppressing the development of specific organs that are not positively involved in sink/source relationships.

Introduction

The domestication of cereal crops such as rice (*Oryza sativa* L.), maize (*Zea mays* L.) and barley (*Hordeum vulgare* L.) began about 10,000 years ago (Doebley et al., 2006). Throughout domestication, plant forms have changed to produce larger edible parts, reduced seed dormancy and loss of shattering (Doebley et al., 2006). Together with these changes some unnecessary organs such as long barbed awns in rice (Hua et al., 2015) and high tillering in maize (Studer et al., 2011) have been lost, resulting in higher grain yields. Grain yield is affected by inflorescence architecture in cereal plants such as rice, maize, barley and wheat (*Triticum aestivum* L.) (Komatsuda et al., 2007; Shomura et al., 2008; Wang et al., 2012; Bommert et al., 2013; Koppolu et al., 2013; Poursarebani et al., 2015). In rice, a model grass plant, several major quantitative trait loci (QTLs) related to grain size such as *GW2*, *qSW5*, *GS5*, *OsSPL16* and *An-1* have been identified (Song et al., 2007; Shomura et al., 2008; Li et al., 2011; Wang et al., 2012; Luo et al., 2013). However, little is known about the genetic basis of grain size in Triticeae species including economically important crops such as wheat, barley and rye (*Secale cereale* L.), except for *TaGW2-A1*, which is a rice *GW2* ortholog in wheat (Simmonds et al., 2016) and *HvDEP1* in barley, an orthologue of rice *Dense and Erect Panicle 1* (Wendt et al., 2016).

As a model organism for plant development in the Triticeae, barley has been useful due to its simple diploid genome and the availability of a large number of different morphological mutants (Druka et al., 2011). In the genus *Hordeum,* including cultivated barley, the inflorescence is called a spike and is composed of three uni-floreted spikelets at each rachis node (Bothmer et al., 1995). When all three spikelets are fertile and produce grains the spike is called six-rowed whereas only central spikelets are fertile in a two-rowed spike. Barley row-type is one of the most important traits affecting yield, especially for grain number, and advancing its molecular understanding is therefore key to yield improvement. Row-type is controlled by at least five loci: *Six-rowed spike 1* (*vrs1* [syn=*HvHox1*]), *Six-rowed spike 2* (*vrs2*), *Six-rowed spike 3* (*vrs3*), *Six-rowed spike 4* (*vrs4*), and *Intermedium spike-c* (*int-c* [syn=*vrs5*]) and all five genes have

been identified (Komatsuda et al., 2007; Ramsay et al., 2011; Koppolu et al., 2013; Bull et al., 2017; van Esse et al., 2017; Youssef et al., 2017). In four of these (*Vrs1*, *Vrs2*, *Vrs4* and *Vrs5*) wild-type alleles encode transcription factors; homeodomain-leucine zipper class I (HD-Zip I), *SHORT INTERNODES* (*SHI*), *RAMOSA2* (*HvRA2*), and *TEOSINTE BRANCHED1* (*HvTB1*) respectively, and in *Vrs3* the wild-type allele putatively encodes a histone H3K9 demethylase. Loss of function at any of these five genes results in varying levels of fertile lateral spikelet development. Selection of naturally occurring six-rowed alleles of *vrs1* and *int-c* during the domestication of barley positively affected production (Komatsuda et al., 2007; Ramsay et al., 2011).

HD-Zip I class genes like *Vrs1* have evolved through a series of gene duplications and divergence through subfunctionalization or neofunctionalization (Ariel et al., 2007; Sakuma et al., 2013). *HvHox2*, which is a paralog of *Vrs1* (*HvHox1*) is highly conserved among cereals and expressed broadly (Sakuma et al., 2010). So far, a loss-of-function mutant of *HvHox2* has not been reported; thus the gene has been considered as essential for plant development. In contrast, *Vrs1* plays more specific roles in the control of lateral spikelet fertility. *Vrs1* is expressed in the lateral florets, especially in the pistil; at least 58 loss-of function mutants have been described and it therefore seems to be a driving force of genetic variation in barley spike development (Komatsuda et al., 2007; Sakuma et al., 2013).

Deficiens barley plants have a two-rowed spike with extremely rudimental lateral spikelets/florets compared with canonical two-rowed cultivars (Fig. 1A). The *deficiens* phenotype is a naturally occurring variant and potentially an allele of *Vrs1* (Habgood and Chambi, 1984). *Deficiens* barley is endemic in Ethiopia where the unique spike of *labile*, which displays a variable number of fertile spikelets at each rachis node, is also localized (Woodward, 1949; Youssef et al., 2014). Old studies report that *deficiens* barley affects the grain size of central spikelets (Powers, 1936; Habgood and Chambi, 1984). Currently, *deficiens* types dominate United Kingdom winter barley seed production, and hence the

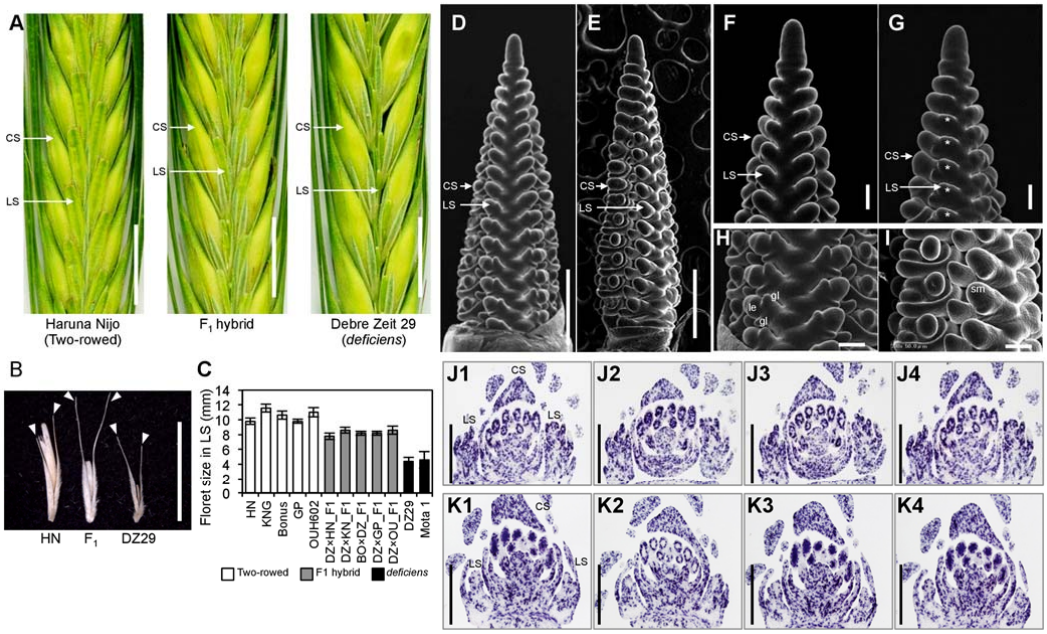

Figure 1. The inflorescence of *deficiens* barley.
(A) The morphology of the inflorescence (from left to right: canonical two-rowed spike cultivar Haruna Nijo, HN; F_1 plant between DZ29 × HN; *deficiens* landrace Debre Zeit 29, DZ29). Central spikelets of DZ29 are larger than those of HN, lateral spikelets of DZ29 are smaller than those of HN. (B) Dissected lateral spikelets. Arrowheads indicate glumes. (C) Variation of floret size in the lateral spikelet. KNG; Kanto Nakate Gold, GP; Golden Promise. (D–I) SEM images of two-rowed cultivar Bonus (D, F, H) and DZ29 (E, G, I) at the awn primordium stage. Asterisks indicate smaller lateral spikelet meristems in DZ29 compared with Bonus. (J and K) *In situ* mRNA hybridization of *histone H4* shows weaker expression in lateral spikelet of DZ29 (K) than that of Bonus (J), while stronger expression in the central spikelet (all the organs) of DZ29 than that of Bonus. Four serial sections are shown in each cultivar. Scale bars = 1 cm in (A), (B), 500 μm in (D), (E), (J), (K) and 100 μm in (F–I). CS: central spikelet; LS: lateral spikelet; gl: glume; le: lemma; sm: spikelet meristem.

area grown, and are increasing in spring barley (Fig. S1). Although breeders recognize the agricultural importance of a perceived yield advantage, the molecular genetic basis of the *deficiens* is not understood. The present study sought to identify the *deficiens* locus through positional cloning and to elucidate the molecular mechanism underlying the extreme suppression of the floral organs of lateral spikelets in barley.

Results

Phenotype of *deficiens* barley

Deficiens barley (Debre Zeit 29; DZ29) shows extremely rudimental lateral floral organs although glume development in its lateral spikelets was comparable to canonical two-rowed barley (Fig. 1A, B). To better understand the *deficiens* phenotype the size of florets in the lateral spikelets was measured at the anthesis stage in field conditions. *Deficiens* barleys produced 4–5 mm long florets in its lateral spikelets, significantly less than the 10-12 mm lengths of canonical two-rowed varieties. F_1 plants from five different parental crosses showed an intermediate phenotype (7–9 mm) compared with parental lines (Fig. 1C) showing a low degree of dominance (0.26, 0.17, 0.23, 0.41, and 0.27, for DZ×HN, DZ×KN, BO×DZ, DZ×GP, and DZ×OU, respectively). To investigate the effect of *deficiens* on the development of lateral spikelets (especially florets) in more detail, scanning electron microscopy (SEM) images were compared (Fig. 1D, E). Barley spike development progresses from the basal to apical part gradually (Kirby and Appleyard, 1981; Youssef et al., 2017). At the apical part of spike, when triple spikelet primordia had differentiated, DZ29 showed a smaller lateral spikelet meristem compared with canonical two-rowed type (Fig. 1F, G). In the canonical type, the lateral spikelet meristems have differentiated into glume and lemma primordia at the basal part, whereas those of DZ29 remained as undifferentiated spikelet meristems (Fig. 1H, I) with suppressed floral differentiation (Fig. 1D–I).

To explore cellular activity of the spikelet meristems, RNA *in situ* hybridization was conducted using *histone H4* antisense probe as a marker for active cell division. The results showed that the transcripts of *histone H4* were detectable in both DZ29 and canonical two-rowed barley (Fig. 1J, K, Fig. S2). Related abundance of the transcript in the lateral spikelets compared to that in the central spikelets is less in DZ29 (Fig. 1K) compared with canonical two-rowed barley (Fig. 1J) although the *in situ* hybridization is not quantitative (Fig. S2). Interestingly, stronger signals were observed in the central spikelet of DZ29, suggesting enhanced cell proliferation. These phenotypic data suggest that

deficiens causes an enduring suppression of lateral spikelet primordia from an early stage of spike development onwards.

Map-based identification of *deficiens*

To map the *deficiens* gene five filial generation 2 (F_2) segregating populations were produced and the genotype of each F_2 individual was inferred by F_3 progeny analysis to show that the genotype segregated as a monogenic factor (Fig. S3, Table S1). Depending upon the cross, the gene was localized to within a 1.0–2.6 cM interval on chromosome 2H (Fig. S4). No recombinants with the HvHox1_DraIII marker, defined by a SNP located on the promoter region of *HvHox1* (–359 bp from transcription start site), were observed. The marker order in the region was conserved across all five mapping populations, indicating no chromosomal rearrangements (Fig. S4).

Larger scale F_2 populations of 2,096 and 2,264 individuals for DZ29 × OUH602 and DZ29 × KNG, respectively, refined the *deficiens* locus to a 0.02 cM region flanked by the markers e40m36-1110S_AccIII and BC12348_HhaI (Fig. 2A, Table S2). The interval is composed of 458,564 bp in Morex BAC sequences (accession no. EF067844) (Komatsuda et al., 2007) and only one predicted gene, *Vrs1* (*HvHox1*) encoding a homeodomain leucine zipper class I transcription factor, was found in the interval. A sequence comparison of the *Vrs1* locus (1,580 bp upstream, 1,538 bp coding region, 1,183 bp downstream) among mapping parents revealed a base pair change that encoded a single amino acid substitution (S184G: 184 serine to glycine) at the Serine-rich motif of the C-terminal region (CTR) specific to DZ29 (renamed *Vrs1.t1* allele, Fig. S5). Resequencing *Vrs1* in 48 NW European winter barley cultivars (24 canonical two-rowed and 24 *deficiens*) revealed that all 24 accessions of the *deficiens* type shared the *Vrs1.t1* allele (Table S3). Furthermore, according to the NCBI database all sequenced accessions of *deficiens* barley have an identical amino acid sequence as VRS1.t1 (Table S3). This suggests that the amino acid change in this region may be responsible for the *deficiens* phenotype.

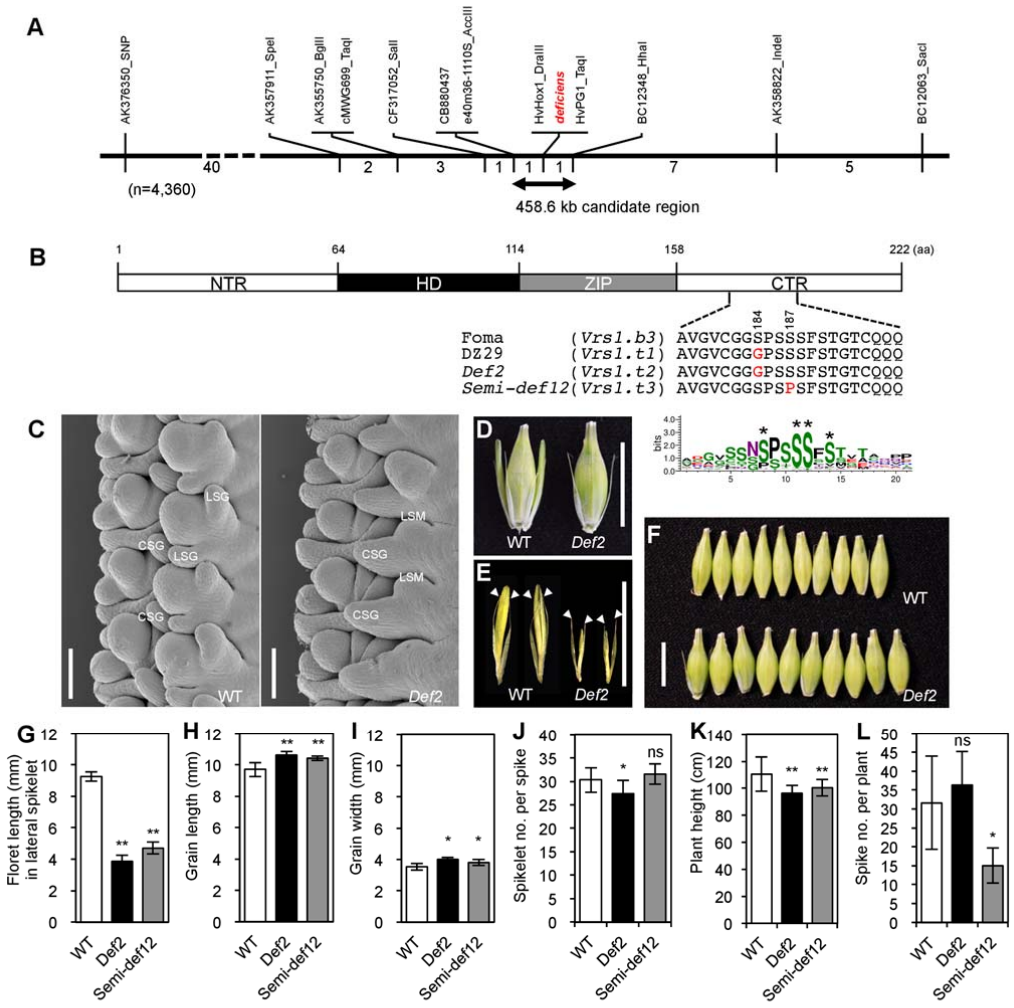

Figure 2. Fine mapping of *deficiens* and analysis of induced mutants.
(A) The genetic map of the *deficiens* region. The numbers beneath the line indicate the number of recombinants recovered. One recombination corresponds to a genetic distance of 0.01 cM. (B) Protein sequence of VRS1. Independent single amino acid substitutions were identified in DZ29 and induced mutants (S184G & S187P) at C-terminal region (CTR), respectively. The Ser-rich motif is highly conserved among plant HD-Zip I proteins (n = 87). Asterisks indicate putative phosphorylation sites. NTR: N-terminal region; HD: homeodomain; ZIP: leucine zipper. (C) SEM images of Foma (WT) and *Deficiens 2* (*Def2*) mutant at awn primordium stage. In *Def2* the lateral spikelet meristem is highly suppressed. CSG: central spikelet glume; LSG: lateral spikelet glume; LSM: lateral spikelet meristem. (D) Triple spikelet morphology. (E) Dissected lateral spikelets. Arrowheads indicate glumes. (F) *Def2* mutant shows bigger grain size. (G–L) Phenotypic comparison between Foma (WT) and *deficiens* mutants. (G) floret length in the lateral spikelet, (H) grain length, (I) grain width, (J) central spikelet number per spike, (K) plant height, (L) spike number per plant. Data are given as mean ± sd (n = 20). **, *, Means are significantly different at the 1 and 5% probability levels, respectively; ns, not significantly different. Scale bars show 100 μm in (C) and 1 cm in (D–F).

To test this hypothesis, seven *Deficiens* mutants and 35 *Semi-deficiens* mutants derived from canonical two-rowed cultivars (Foma, Bonus, Kristina and Lina) were phenotyped and re-sequenced (Table S4). Two of the 42 induced mutants,

Deficiens 2 (*Def2,* named *Vrs1.t2* allele*)* and *Semi-deficiens 12* (*Semi-def12*, named *Vrs1.t3* allele*)* derived from cv. Foma, showed a clear *deficiens* spike phenotype with rudimentary lateral spikelets (Fig. 2C–E, G, Fig. S6). These two mutants had independent single amino acid substitutions, S184G and S187P respectively, at the CTR (Fig. 2B, Table S4) suggesting that a single amino acid substitution of serine in the Ser-rich motif was responsible for the *deficiens* phenotype (Fig. S7). The remaining 40 lines did not show a *deficiens* phenotype nor a sequence variant, suggesting that either the character has been lost from the genebank stocks or the phenotype was environmentally controlled and not genetic (Table S4). SEM images of *Def2* showed that lateral spikelet meristems seem to arise from the central spikelet glume (Fig. 2C and Fig. S8). Interestingly, the *Def2* mutant shares an identical SNP (A > G at position +775 relative to the start codon) with *deficiens* cultivars including DZ29. To date, *Vrs1* has been resequenced in more than four hundred lines but neither the nucleotide sequences (nor haplotypes) of *Def2* nor *Semi-def12* mutants were found in any databases suggesting this change is novel, originating from a mutation in cv. Foma. Interestingly, the Ser-rich motif around the position of 184 and 187 was highly conserved amongst plant HD-Zip I class transcription factors (*n* = 87; 22 species in Embryophyta, Fig. 2B, Table S5 and Supplementary Data 1). Since serine is known as a potential phosphorylation site, we examined the possibility of phosphorylation of serine using the program NetPhos 3.1. The results showed that the serine residues at CTR in canonical two-rowed genotypes are predicted targets of phosphorylation, whereas *deficiens* alleles showed lower potential than the threshold (Fig. S9) suggesting an altered phosphorylation status in *deficiens* lines.

To understand the phenotypic effect of the *deficiens* mutations, several agronomic traits were compared between Foma (WT) and its derived *Def2* and *Semi-def12* mutants. Both *deficiens* mutants showed about a 10% increase in grain length and width compared with WT (Fig. 2F, H, I). Moreover, both *deficiens* mutants showed reduced plant height compared with WT (Fig. 2K).

The significant difference of spikelet number per spike was only observed in *Def2* and spike number per plant in *Semi-def12*.

The phenotypic effect of *deficiens* (*Vrs1.t1*) was then assessed in the collection of 48 contemporary North West European winter barley cultivars (24 canonical two-rowed and 24 *deficiens*) referred to previously for yield and yield related traits (Figure 3). The *deficiens* lines have a significantly higher thousand grain weight (56.5 g) than the canonical two-rowed lines (53.5 g) (t (23), p=0.002). In this set of germplasm the number of spikes per m^2 was slightly, but not significantly, lower for the *deficiens* lines (629.8) compared to two-rowed lines (665.3) (t (23), p=0.051) and *deficiens* does not significantly influence yield (*deficiens* = 7.2 t/ha, two-rowed = 7.1 t/ha) or grains per spike (*deficiens* = 20.5, two-rowed = 20.1).

Expression pattern of *Vrs1* in *Deficiens 2* mutant

Quantifying mRNA levels at different developmental stages revealed no significant differences in *Vrs1* transcript abundance at the beginning of spike development (from double-ridge to lemma primordium stage) between WT and *Def2* mutant. This indicates that *Vrs1* is normally regulated at this stage (Fig. 4A). After the lemma primordium stage, however, the *Vrs1* transcript level was significantly decreased in both *Def2* (Fig. 4A) and DZ29 (Fig. S10). To determine whether or not the reduced *Vrs1* transcript level was affected by the upstream regulator *Vrs4*, we examined *Vrs4* expression levels but found no significant difference between WT and *Def2* (Fig. 4B). The expression pattern of *HvHox2*, which is a paralog of *Vrs1*, was also not significantly different (Fig. 4C).

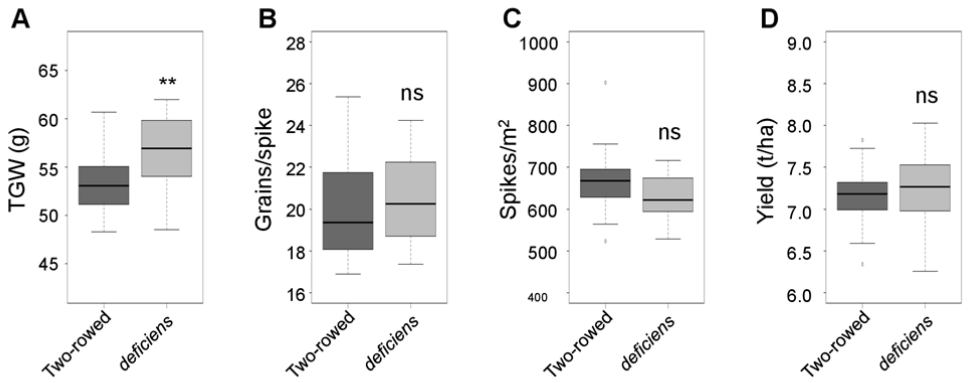

Figure 3. Mean effect of *deficiens* allele for grain yield.
(A) Thousand grain weight (TGW). (B) Grains per spike. (C) Spikes per m2. (D) Yield (t/ha). Data were obtained from 5 site/seasons field trial of 24 two-rowed and 24 *deficiens* NW European cultivars grown under a winter feed barley regime (200kg total N/ha).
**, *, Means are significantly different at the 1 and 5% probability levels, respectively; ns, not significantly different.

The hypothesis that a different tissue localization of the *Vrs1* transcript could also be responsible for the *deficiens* phenotype guided us to test this possibility by RNA *in situ* hybridization using *Vrs1* specific probes (Sakuma et al., 2013). *Vrs1* mRNA was clearly localized to the lateral spikelet meristem in both *deficiens* and canonical two-rowed barley at the glume primordium stage and the white anther stage respectively (Fig. 4D–G). These qRT-PCR and *in situ* hybridization data suggest that the amino acid alterations of VRS1, but not mRNA differences, were responsible for the *deficiens* phenotype.

Genome-wide transcriptional regulation by *Deficiens 2*

To identify the transcriptional changes underlying the morphological differences during immature spike development, RNA-seq was conducted using developing spikes at the awn primordium stage of both WT and the *Def2* mutant. Based on SEM analysis at this developmental stage the meristem of the lateral spikelet remains undifferentiated in *Def2* (Fig. 2C). We reasoned that such a clear morphological difference should enable the detection of differentially expressed genes (DEGs). RNA-seq analysis revealed 1,256 DEGs (q-value < 0.05, p-value < 0.05 and $\log_2$ fold change > 1 or < -1); comprised of 860 (68.5%) upregulated and 396 (31.5%) downregulated genes in the *Def2* mutant relative to WT (Fig. 5A, Table S6). Gene Ontology (GO) enrichment analysis of the differentially

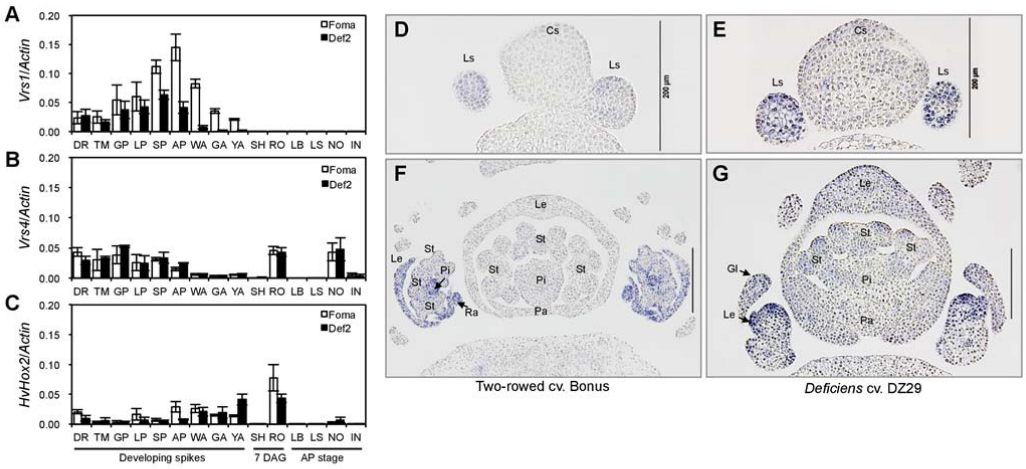

Figure 4. Transcript pattern of *Vrs1*, *Vrs4* and *HvHox2*.
(A–C) Quantitative real time PCR comparison of *Vrs1* (A), *Vrs4* (B), and *HvHox2* (C) in selected tissues between Foma and *Deficiens 2* mutant. DR: double ridge stage; TM: triple-mound stage; GP: glume primordium stage; LP: lemma primordium stage; SP: stamen primordium stage; AP: awn primordium stage; WA: white anther stage; GA: green anther stage; YA: yellow anther stage; SH: shoots; RO: roots; LB: leaf blade; LS: leaf sheath; NO: nodes; IN: internodes. (D–G) *In situ* mRNA hybridization with *Vrs1* antisense probe. *Vrs1* is predominantly expressed in the lateral spikelet meristem at the glume primordium stage of Bonus (*Vrs1.b3*, D) and DZ29 (*Vrs1.t1*, E). At the white anther stage, *Vrs1* is mainly localized in lemma, pistil and rachilla in Bonus (F) whereas *Vrs1* is expressed in floret meristem and lemma primordia in DZ29 (G). Cs: central spikelet; Ls: lateral spikelet; Gl: glume; Le: lemma; St: stamen; Pi: pistil; Ra: rachilla. Scale bars show 200 μm.

expressed genes revealed that genes upregulated in *Def2* (Supplementary Data 3) were highly enriched for molecular functions related to DNA binding (GO:0003677, P = 5.80e–11) and glutathione transferase activity (GO:0004364, P = 1.20e–10) (Fig. 5B and Table S7). The DNA binding class includes 78 *histone* superfamily genes, indicating promotion of cell proliferation in *Def2*; this might be related with the fact that grain size was increased in *deficiens* alleles (Fig. 2). In contrast, genes downregulated in *Def2* were not generally enriched for any GO terms (Supplementary Data 4), although we did detect downregulation of cytokinin oxidase/dehydrogenase (*CKX*) genes, which are involved in cytokinin degradation (HORVU3Hr1G105360 and HORVU3Hr1G075920). Reduced expression of *CKX* gene generally causes cytokinin accumulation and more meristem activity (Ashikari et al., 2005; Han et al., 2014). The reduction of *Vrs1* (HORVU2Hr1G092290) expression level in *Def2* was also detected in RNA-seq analysis, which validates the findings of the qRT-PCR experiment.

The origin of *deficiens* barley

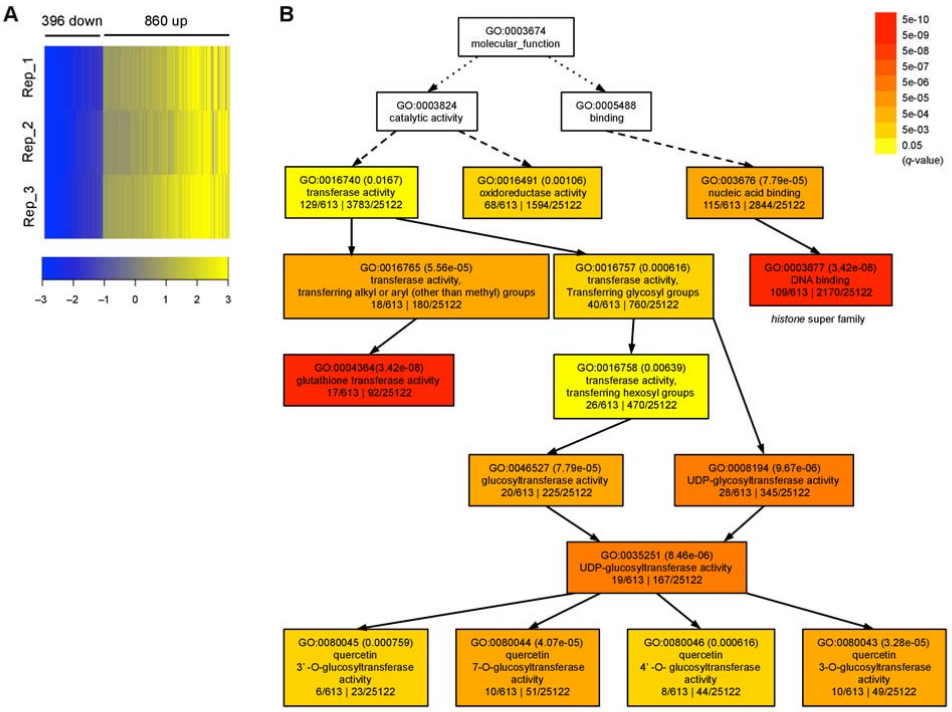

Figure 5. Transcriptional regulation by the *Deficiens* mutation.
(A) Heat map of differentially expressed genes (DEG) in *Def2* mutant (*Vrs1.t2* allele) compared with Foma. (B) Hierarchical tree graph of Gene Ontology (GO) enrichment for the up-regulated DEG. Numbers in parentheses indicate *q*-values. Solid, dashed, and dotted lines represent two, one and zero enriched terms at both ends connected by the line, respectively.

The *deficiens* landraces are predominantly distributed in Ethiopia and a monophyletic origin of *deficiens* is hypothesized. To infer the origin of the *deficiens* allele (*Vrs1.t1*), we reanalyzed the sequence of the *Vrs1* region (2062 bp length) from 321 domesticated barley and 136 wild barley accessions (Saisho et al., 2009). Haplotype network analysis revealed that *Vrs1.t1* allele (Haplotype 10) was derived from the *Vrs1.b2* allele (Haplotype 6) through a single nucleotide substitution in exon 3, which resulted in the S184G amino acid substitution implying that *Vrs1.b2* was the immediate ancestral allele of *Vrs1.t1*

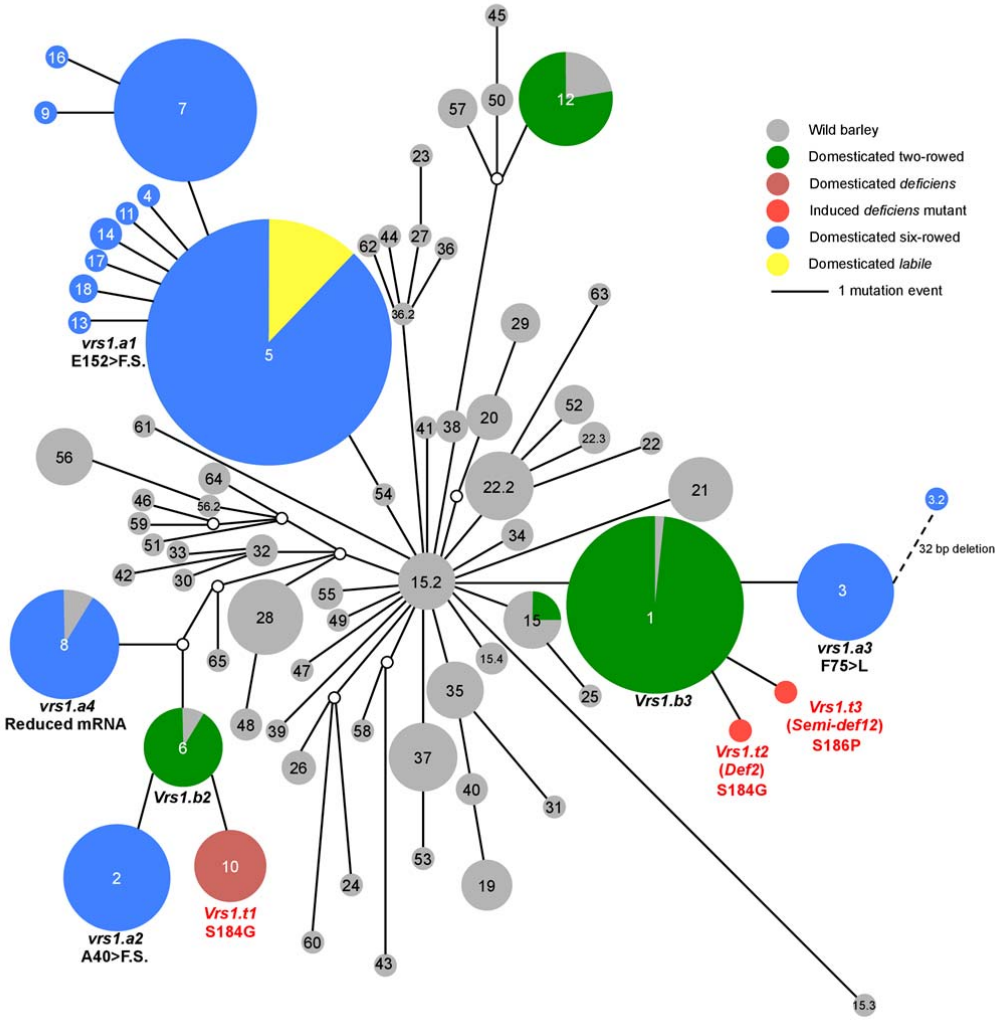

Figure 6. Haplotype network analysis of *Vrs1*.
Median-Joining network was constructed from the haplotypes based on the re-sequencing of 321 domesticated and 136 wild barley accessions. Circle sizes correspond to the frequency of individual haplotypes. Lines represent genetic distances between haplotypes. Open circle indicate inferred intermediate haplotypes.

(Fig. 6). Landraces carrying Haplotype 6 are predominantly distributed in North Africa (Table S8). Interestingly, the *vrs1.a2* allele (Haplotype 2), a six-rowed spike, was also derived from the same *Vrs1.b2* allele through a single nucleotide insertion in exon 2, which results in a frame shift of the amino acid sequence (A40>F.S.; Fig. 6). Landraces carrying the *vrs1.a2* allele are distributed throughout North Africa, South Europe and North America. Both *Vrs1.t1* (contributing grain size) and *vrs1.a2* (contributing grain number) alleles originated from North Africa. While *Vrs1.t1* was endemic to Ethiopia, *vrs1.a2*

expanded to North Africa and South Europe. We consider it remarkable that the opposite yield components, grain size and grain number, are controlled by different alleles derived from the same *Vrs1* gene and ancient people selected both mutations independently.

Considering contemporary North West European winter cultivars, we identified 3 SNPs in the CDS of *Vrs1* leading to amino acid substitutions, D8G, D26E, S184G, and one SNP in the 2nd intron of this gene (Table S3). These 4 SNPs are in complete linkage disequilibrium, and the two haplotypes segregate based on the canonical two-rowed/*deficiens* phenotype of each cultivar. All of the *deficiens* lines contain the *Vrs1.t1* allele, shared with accessions originating from Ethiopia (Table S3).

Discussion

A single amino acid substitution in VRS1 is responsible for *deficiens* barley

Deficiens has previously been considered one of the multiple alleles at the *Vrs1* locus (Woodward, 1947), but its causal mutation had not been discovered. In the present study, detailed phenotypic observation confirmed that *deficiens* alleles extremely suppress floret development in the lateral spikelets. Our genetic analysis strongly suggests that a point mutation leading to a single amino acid substitution at the conserved Ser-rich motif in C-terminal region (Motif 4) of VRS1 is responsible for the developmental fate of the lateral spikelet. This Ser-rich motif is highly conserved among plant HD-Zip I class transcription factors but its biological function(s) remained elusive. Next to threonine and tyrosine, serine is an amino acid residue commonly phosphorylated by several kinases in eukaryotes. Large-scale phosphorylation mapping analysis revealed 85% of serine residues in Arabidopsis to be phosphoserine (Sugiyama et al., 2008). Phosphorylation influences the stability and activity of transcription factors (Hunter and Karin, 1992; Fujiwara et al., 2008). When phosphorylated, CONSTANS (CO), a flowering promoter encoding a B-box zinc finger, influences the rate of turnover of the CO protein via activity of the CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) ubiquitin ligase (Sarid-Krebs et al., 2015). Phosphorylated CONSTANS is the preferred substrate for degradation. In this study, almost all serine residues in the Ser-rich motif (Motif 4) were predicted as phosphoserine in the canonical two-rowed allele of *Vrs1* but not in its *deficiens* alleles because of a single amino acid substitution in Motif 4. From our results, we hypothesize that hypo-phosphorylated VRS1 in *deficiens* remains functional for longer during a critical stage in plant development, resulting in stronger suppression of floret development in the lateral spikelets. The expression analysis revealed that the *deficiens* phenotype is not due to an overexpression of *Vrs1* but most likely acts at the protein level (Fig. 4). Reduced *Vrs1* transcript levels from the stamen primordium stage onwards are probably due to the much smaller size of the lateral spikelets, compared with WT, noting that *Vrs1* mRNA is predominantly localized in the lateral spikelets (Fig. 4D–G).

Since *Vrs1* evolved through gene duplication and neofunctionalization, the mutation frequency is notably higher than in its paralog *HvHox2* (Sakuma et al., 2010; Sakuma et al., 2013). Six of the 58 *vrs1* mutants identified so far have a mutation at the C-terminal region (Motif 4 end and motif 10 in this study), which is considered as a transcriptional activation domain (Komatsuda et al., 2007; Arce et al., 2011). This region is in the vicinity of the phosphorylation region, and the mutation sites were C194>S (*Int-d.11*), Q196>stop (*Int-d.36*), Q197>stop (*Int-d.40*) and A203>frame shift (*hex-v.27*, *hex-v.30* and *hex-v.31*). Previous yeast two hybrid assays showed that the C-terminal region (Motifs 4 and 10 in this study) of VRS1 functions as a transcriptional activator (Sakuma et al., 2013). Therefore, the six *vrs1* mutants mentioned above could each be considered to represent a loss of transactivation activity as a transcription factor. By identifying its role in the suppression of floret primordia, the present study is the first to uncover a biological function of the Ser-rich motif in HD-Zip I transcription factors.

Higher yield potential of *deficiens* barley

Deficiens is now the major spike type of winter barleys grown in NW Europe, and is increasing in the European spring barley crop (Fig. S1). We therefore hypothesize that *deficiens* spike types have replaced canonical two-rowed types due to a higher grain yield potential associated with the altered spike architecture. Our results strongly support this hypothesis since two independent mutant accessions showed enlarged grains, which is positively related to higher grain yield. It is hard to think of why *deficiens* resulted in reduced height; though a possible explanation is a change in resource partitioning or additional mutations at other loci that reduce height. A previous study showed that the sterile lateral spikelets of canonical two-rowed barleys do not contribute to photosynthesis suggesting that there is potential conflict between the central and lateral spikelets for assimilate supply (Habgood and Chambi, 1984). The notion that *vrs1* loss-of-function mutants produce more but smaller-grains compared to WT supports this model. Improvement of grain size could therefore

be achieved by suppressing the development of specific organs (in this study; lateral spikelets), which do not contribute to yield potential.

Our RNA-seq analysis revealed that a number of *histone* genes were upregulated in the *Def2* mutant suggesting active cell proliferation in the central spikelet which in turn may result in bigger grain size. This hypothesis is partially supported by the *in situ* hybridization analysis with *histone H4* probe showing higher signals in the central spikelet and lower signals in the lateral spikelet of *deficiens* barley (Fig. 1J, K). Since lateral florets were strongly suppressed at the awn primordium stage, the upregulation of *histone* genes would be a central spikelet event and an indirect effect through the suppression of lateral floret development. In rice, higher expression of *GW6a* enhances grain weight and yield by increasing acetylation levels of histone H4 and upregulation of cell cycle genes (Song et al., 2015). Interestingly, the rare allele elevating *GW6a* expression has escaped human selection during rice domestication and modern breeding. Enlarged grain size and thus TGW in *deficiens* barley, as observed in both mutant germplasm and NW European cultivars, could be analogous to the *GW6a* event in rice, albeit at an independent locus. However, the precise mechanism by which upregulating cell cycle genes pre-fertilization leads to increases in grain size remains to be established.

Haplotype network analysis revealed that both *Vrs1.t1* and *vrs1.a2* alleles originated from a point mutation of the *Vrs1.b2* allele. Based on geographical data, it stands to reason that mutation of *Vrs1.b2* to *Vrs1.t1* occurred in Ethiopia and was selected there while *vrs1.a2* was selected in North-west Africa and then distributed to Southern Europe. Remarkably, two different alleles with opposing phenotypic effects, enlarged grain size and increased grain number, were selected independently during barley domestication, presumably due to different selective forces for these two competing components of yield. *Vrs1.t1* can be considered a rare allele, and its utilization has increased in NW European winter and spring barley only recently.

In summary, we identified that the *deficiens* spike architecture of cultivated barley is the result of a single amino acid substitution in a conserved motif of the major row-type protein VRS1. The substitution leads to a repression of the development of the lateral florets on the barley spike. Our phenotypic analysis suggested that *deficiens* alleles (*Vrs1.t1*, *.t2*, and *.t3*) ultimately contribute to increasing the size of grain in the central spikelets. Haplotype analysis indicates that the natural *deficiens Vrs1.t1* allele originated from a point mutation in the two-rowed allele *Vrs1.b2*, an event that most probably occurred in Ethiopia. Unlocking the precise molecular cause of the *deficiens* spike morphology will enable direct selection for higher yielding two-rowed barley varieties and enhance our biological understanding of the molecular mechanisms underpinning the elaboration of barley inflorescence development.

Material and Methods

Plant materials

The *deficiens* barley (*Hordeum vulgare* var. *deficiens*) Debre Zeit 29 (DZ29, SV043) and Mota 1 (SV039) together with wild barley line OUH602 (*H. vulgare* ssp. *spontaneum*) were obtained from the Institute of Plant Science and Resources, Okayama University, Kurashiki, Okayama, Japan. Canonical two-rowed cultivars (cv.), Haruna Nijo (HN), Kanto Nakate Gold (KNG), Bonus (BO), Foma, Golden Promise (GP) were obtained from National Institute of Agrobiological Sciences (NIAS), Tsukuba, Japan. The *Deficiens* and *Semi-deficiens* mutants listed in Table S4 were obtained from the Nordic gene bank (NordGen, Alnarp, Sweden). 24 *deficiens* and 24 canonical two-rowed cultivars grown under a feed nitrogen regime were analyzed in the current paper. They were first placed on the UK National List between 1992 – 2012, and 1992 – 2007 respectively, representing as far as possible a balanced dataset, and focusing on contemporary, relevant germplasm. The field trials were grown over 5 site/season combinations in Scotland, England, and Germany for harvest years 2010 and 2011 with known available soil mineral nitrogen content. Each trial was supplemented with three fertilizer regimes in a split plot design with partial replication; no nitrogen application, a malting nitrogen application and a

feed nitrogen application. Plots were sown at a constant density and kept free from disease with a standard fungicide regime. Grab samples were taken from each plot at physiological maturity and used to estimate spikes/m2 and grains/spike. The plots were harvested, grain dried to a constant moisture content, weighed, cleaned and a sub-sample used to estimate thousand grain weight using a Marvin digital seed analyzer (GTA Sensorik). The weighed grain from each plot together with the plot area was used to estimate grain yield in t/ha. Not all lines were in trial for two years and so Best Linear Unbiased Predictors (BLUPs) were calculated for each genotype and trait with the nitrogen treatment as a fixed effect and all others, including the interaction of line with fertilizer, as random effects. BLUPs were generated using the REML directive in Genstat 14 (http://www.vsni.co.uk/software/genstat/). All lines were previously genotyped with the barley iSelect 9K SNP platform (Comadran et al., 2012).

Scanning electron microscopy (SEM) analysis

Immature spikes were fixed in 4% formaldehyde in phosphate buffer. After dehydration and critical point drying, samples were examined in a Hitachi S4100 or a KEYENCE VHX-D510 scanning electron microscope at 5kV acceleration voltage.

Basic mapping

Five F_2 populations were developed by crossing DZ29 with five different two-rowed accessions (KNG, HN, BO, GP and OUH602). F_2 plants were grown until full maturity under field condition at NIAS, Tsukuba, Japan. To construct a basic genetic map, 78 to 96 plants were used from each of the five F_2 populations. The *deficiens* genotype of F_2 plants was determined by progeny testing 24–30 F_3 plants from each F_2 individual in the field of Central Agricultural Experiment Station, Hokkaido, which enabled classification into the three genotypic classes expected for the segregation of a single Mendelian gene. The length of the lateral florets from the centre of at least three spikes per plant were measured after spike emergence was complete. F_2 individuals were genotyped using selected polymorphic DNA markers from the insertion-deletion (Indel),

single nucleotide polymorphism (SNP), cleaved amplified polymorphic sequence (CAPS) and derived CAPS (dCAPS) markers described by Sakuma et al. (2010). The primers are listed in Table S9 and the resulting genotypic data used to construct linkage maps using MAPMAKER/EXP ver.3.0 (Lander et al., 1987).

Fine mapping

The F_2 segregants of DZ29 × KNG and DZ29 × OUH602 were genotyped with two DNA markers AK376350_SNP and BC12063_SNP3 that we found to flank the *deficiens* locus. F_3 progeny in which a recombination had occurred between the two flanking markers were genotyped with *de novo* DNA markers to further delimit the *deficiens* locus (Table S9).

Resequencing *Vrs1* in NW European winter barley cultivars

Genomic DNA was extracted using bench-grown 2 week old leaf tissue with a QIAamp 96 DNA kit on a QIAcube HT automated DNA extraction platform. (Qiagen, Hilden, Germany). DNA quality and quantity was assessed using a NanoDrop 2000 (Thermo Scientific, Massachusetts, USA). PCR amplifications and Sanger sequencing were carried out as described previously (Houston et al., 2012). DNA sequence trimming and alignments were performed using Geneious (vR10, Biomatters Limited, Auckland, NZ). Details of primers used are provided in Table S9.

Peptide motif analysis and functional prediction

To obtain VRS1 homologues from several plant species a BLASTP search (query: BAF43315) was made against the Phytozome v.12.0 (https://phytozome.jgi.doe.gov/pz/portal.html) and IPK Barley BLAST Server (http://webblast.ipk-gatersleben.de/barley_ibsc/). Peptide sequences with a threshold of *E*-value < 1e–10 were retrieved and 87 peptide sequences (Supplementary Data 1) including VRS1 and HvHOX2 (BAI49294) were used for motif analysis. MEME database was used for discovering peptide motifs (Bailey et al., 2006). Prediction of potential phosphorylation amino acid sites was performed using NetPhos 3.1 (Blom et al., 1999; Blom et al., 2004).

532

533 ***In situ* RNA hybridization analysis**

534 Immature spikes at the glume primordium stage and the white anther stage
535 (Kirby and Appleyard, 1981), from field-grown plants were freshly sampled and
536 fixed in 4% paraformaldehyde and 0.25% glutaraldehyde in 0.05M
537 sodium-phosphate buffer pH 7.2. RNA probes of *Vrs1* and *histone H4* were
538 developed as in Sakuma et al. (2013). Primers are listed in Table S9. *In situ*
539 hybridization was conducted as described by Komatsuda et al. (2007).

540

541 **RNA extraction and quantitative real-time PCR**

542 Immature spikes were developmentally staged from double ridge stage to yellow
543 anther stage under a stereoscopic microscope (Kirby and Appleyard, 1981).
544 Total RNA was extracted from immature spikes, seedling shoots and roots 7
545 days after germination, leaf blades, leaf sheaths, nodes and internodes at awn
546 primordium stage using TRIzol (Invitrogen, Carlsbad, CA, USA). RNA was
547 quantified using a NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA,
548 USA). To remove genomic DNA contamination, RNA was treated with
549 RNase-free DNase (Takara Bio, Tokyo, Japan). First-strand cDNA was
550 synthesized with SuperScript III (Invitrogen) and first-strand cDNA derived from
551 20 ng RNA was used as PCR template. Transcript levels of each gene were
552 measured by qRT-PCR using an ABI Prism 7900HT sequence detection system
553 (Applied Biosystems, Foster, CA, USA) and THUNDERBIRD SYBR qPCR Mix
554 Kit (Toyobo, Osaka, Japan) according to the manufacturers' protocols. Primers
555 used for qRT-PCR are listed in Table S9. qRT-PCR analysis was performed at
556 least twice for each sample with biological replicates of at least three
557 independent RNA extractions per sample. Barley *Actin* gene (accession
558 DN182500) was used to normalize the RNA level for each sample.

559

560 **RNA-seq**

561 Total RNA was extracted from immature spikes at the awn primordium stage of
562 cv. Foma and a *Deficiens 2* mutant (NGB115174). Total RNA was measured
563 using an Agilent 2100 Bioanalyzer (Agilent Technologies, Technologies, PaloAlto,

CA, USA) and used for the construction of sequencing libraries. Strand-specific RNA libraries were prepared using the TruSeq Stranded mRNA Sample Prep Kit (Illumina) following the instructions in the TruSeq Stranded mRNA Sample Preparation Guide Rev.E (Illumina). High-throughput sequencing was conducted using a HiSeq 2500 (Table S10, Illumina, San Diego, CA, USA). Sequences were aligned to barley pseudomolecules (Mascher et al. (2017): 160404_barley_pseudomolecules_masked.fasta) with TopHAT2 (Kim et al., 2013). Gene expression was estimated as read counts for each gene locus by featureCounts (Liao et al., 2013) using the gene annotation file Hv_IBSC_PGSB_r1_HighConf.gtf. Expression levels were normalized by TMM method and *p*-values were calculated by an exact negative binomial test along with the gene-specific variations estimated by empirical Bayes method in edgeR (Robinson et al., 2010). The Benjamini-Hochberg method was applied on the *p*-values to calculate *q*-values and to control the false discovery rate (FDR). Differentially expressed genes (DEGs) were defined as *q*-value < 0.05, *p*-value < 0.05 and $\log_2$ fold change > 1 or < -1. The sequence data have been submitted to the DDBJ Sequence Read Archive under accession number PRJDB4754.

Gene ontology (GO) term enrichment analysis

GO terms of barley genes were assigned using Arabidopsis GO terms as follows. *Hordeum vulgare* protein sequences (160517_Hv_IBSC_PGSB_r1_proteins_HighConf_REPR_annotation.fasta) were compared with the peptide sequence in *Arabidopsis thaliana* (TAIR10; (Berardini et al., 2015)) using BLAST software (blastp) with a threshold of *E*-value < 1e–20; (Altschul et al., 1990). GO terms with most similar sequence in *Arabidopsis* were used as the GO terms of corresponding barley genes. 25,122 barley genes were assigned with more than one GO term (Supplementary Data 2). GO term enrichment analysis of up- (Supplementary Data 3) and down-regulated genes (Supplementary Data 4) were performed separately using agriGO (Du et al., 2010) with Fisher's exact test and Benjamini, Hochberg, and Yekutieli method to control FDR.

Haplotype analysis

A haplotype analysis was performed based on re-sequencing data of the *Vrs1* locus (2062 bp) from 321 domesticated barley accessions and 136 wild barley accessions (Table S8; Saisho et al., 2009). Sequence alignments were performed with ClustalW using MEGA7 software (Kumar et al., 2016). A Median-Joining network (Bandelt et al., 1999) was constructed using the software programs DNA Alignment v1.3.3.2 and Network v5.0.0.1 (Fluxus Technology Ltd., Clare, Suffolk, UK) with default parameters (Epsilon=0; Frequency>1 criterion=inactive; The ratio of transversion:transition=1:1; Criterion=Connection cost; External rooting=inactive; MJ square option=inactive).

Accession numbers

The *Vrs1* sequence data have been deposited in DNA Data Bank of Japan (DDBJ) nucleotide core database under accession numbers LC216328 – LC216332, except for the NW European cultivars listed in Table S3 which have been deposited in NCBI under accession numbers MF776900 – MF776946. . The RNA-seq data have been submitted to the DDBJ Sequence Read Archive under accession number PRJDB4754.

Acknowledgments

We thank Daisuke Saisho (Okayama University), Naoki Sentoku (NIAS) and Hiroyuki Tsuji (Yokohama City University) for their help and advice. We are grateful to Harumi Koyama and Mari Sakuma (NIAS), Corinna Trautewig (IPK), and Malcolm Macaulay (JHI) for their excellent technical assistance. We thank Peter Werner, Guillaume Barol-Baron, and David Harrap of KWS UK Ltd for allowing us to use some of their data from the trials and we also acknowledge the assistance of Hazel Bull, Richard Keith, and Chris Warden of JHI.

Competing financial interests:

The authors declare no competing financial interests.

628 **Figure legends**

629 **Figure 1.** The inflorescence of *deficiens* barley.

630 (A) The morphology of the inflorescence (from left to right: canonical two-rowed
631 spike cultivar Haruna Nijo, HN; F_1 plant between DZ29 × HN; *deficiens* landrace
632 Debre Zeit 29, DZ29). Central spikelets of DZ29 are larger than those of HN,
633 lateral spikelets of DZ29 are smaller than those of HN. (B) Dissected lateral
634 spikelets. Arrowheads indicate glumes. (C) Variation of floret size in the lateral
635 spikelet. KNG; Kanto Nakate Gold, GP; Golden Promise. (D–I) SEM images of
636 two-rowed cultivar Bonus (D, F, H) and DZ29 (E, G, I) at the awn primordium
637 stage. Asterisks indicate smaller lateral spikelet meristems in DZ29 compared
638 with Bonus. (J and K) *In situ* mRNA hybridization of *histone H4* shows weaker
639 expression in lateral spikelet of DZ29 (K) than that of Bonus (J), while stronger
640 expression in the central spikelet (all the organs) of DZ29 than that of Bonus.
641 Four serial sections are shown in each cultivar. Scale bars = 1 cm in (A), (B), 500
642 μm in (D), (E), (J), (K) and 100 μm in (F–I). CS: central spikelet; LS: lateral
643 spikelet; gl: glume; le: lemma; sm: spikelet meristem.

644

645 **Figure 2.** Fine mapping of *deficiens* and analysis of induced mutants.

646 (A) The genetic map of the *deficiens* region. The numbers beneath the line
647 indicate the number of recombinants recovered. One recombination
648 corresponds to a genetic distance of 0.01 cM. (B) Protein sequence of VRS1.
649 Independent single amino acid substitutions were identified in DZ29 and induced
650 mutants (S184G & S187P) at C-terminal region (CTR), respectively. The
651 Ser-rich motif is highly conserved among plant HD-Zip I proteins (*n* = 87).
652 Asterisks indicate putative phosphorylation sites. NTR: N-terminal region; HD:
653 homeodomain; ZIP: leucine zipper. (C) SEM images of Foma (WT) and
654 *Deficiens 2* (*Def2*) mutant at awn primordium stage. In *Def2* the lateral spikelet
655 meristem is highly suppressed. CSG: central spikelet glume; LSG: lateral
656 spikelet glume; LSM: lateral spikelet meristem. (D) Triple spikelet morphology.
657 (E) Dissected lateral spikelets. Arrowheads indicate glumes. (F) *Def2* mutant
658 shows bigger grain size. (G–L) Phenotypic comparison between Foma (WT) and
659 *deficiens* mutants. (G) Floret length in the lateral spikelet. (H) Grain length. (I)

Grain width. (J) Central spikelet number per spike. (K) Plant height (L) Spike number per plant. Data are given as mean ± sd (*n* = 20). **, *, Means are significantly different at the 1 and 5% probability levels, respectively; ns, not significantly different. Scale bars show 100 µm in (C) and 1 cm in (D–F).

Figure 3. Mean effect of *deficiens* allele for grain yield.
(A) Thousand grain weight (TGW). (B) Grains per ear. (C) Spikes per m2. (D) Yield (t/ha). Data were obtained from 5 site/seasons field trial of 24 two-rowed and 24 *deficiens* NW European cultivars grown under a winter feed barley regime (200kg total N/ha). **, *, Means are significantly different at the 1 and 5% probability levels, respectively; ns, not significantly different.

Figure 4. Transcript pattern of *Vrs1*, *Vrs4* and *HvHox2*.
(A–C) Quantitative real time PCR comparison of *Vrs1* (A), *Vrs4* (B), and *HvHox2* (C) in selected tissues between Foma and *Deficiens 2* mutant. DR: double ridge stage; TM: triple-mound stage; GP: glume primordium stage; LP: lemma primordium stage; SP: stamen primordium stage; AP: awn primordium stage; WA: white anther stage; GA: green anther stage; YA: yellow anther stage; SH: shoots; RO: roots; LB: leaf blade; LS: leaf sheath; NO: nodes; IN: internodes. (D–G) *In situ* mRNA hybridization with *Vrs1* antisense probe. *Vrs1* is predominantly expressed in the lateral spikelet meristem at the glume primordium stage of Bonus (*Vrs1.b3*, D) and DZ29 (*Vrs1.t1*, E). At the white anther stage, *Vrs1* is mainly localized in lemma, pistil and rachilla in Bonus (F) whereas *Vrs1* is expressed in floret meristem and lemma primordia in DZ29 (G). Cs: central spikelet; Ls: lateral spikelet; Gl: glume; Le: lemma; St: stamen; Pi: pistil; Ra: rachilla. Scale bars show 200 µm.

Figure 5. Transcriptional regulation by the *Deficiens* mutation.
(A) Heat map of differentially expressed genes (DEG) in *Def2* mutant (*Vrs1.t2* allele) compared with Foma. (B) Hierarchical tree graph of Gene Ontology (GO) enrichment for the up-regulated DEG. Numbers in parentheses indicate *q*-values. Solid, dashed, and dotted lines represent two, one and zero enriched terms at

692 both ends connected by the line, respectively.
693
694 **Figure 6.** Haplotype network analysis of *Vrs1*.
695 Median-Joining network was constructed from the haplotypes based on the
696 re-sequencing of 321 domesticated and 136 wild barley accessions. Circle sizes
697 correspond to the frequency of individual haplotypes. Lines represent genetic
698 distances between haplotypes. Open circle indicate inferred intermediate
699 haplotypes.
700
701 **Supplemental Material**
702 **Table S1.** Segregation analysis of deficiens in five F2 mapping populations.
703 **Table S2.** The mapping population used for fine mapping of *deficiens.*
704 **Table S3.** *Vrs1* sequence information of *deficiens* barley.
705 **Table S4.** List of *Deficiens* and *Semi-deficiens* mutants with description of
706 mutational events in *Vrs1.*
707 **Table S5.** Motif analysis of HD-Zip I transcription factor in plants (n = 87).
708 **Table S6.** Differentially regulated genes in *Deficiens2* mutants compared to
709 wild-type.
710 **Table S7.** Gene Ontology (GO) enrichment analysis for up-regulated
711 differentially expressed genes in *Deficiens 2* mutant.
712 **Table S8.** Haplotype information of 321 domesticated and 136 wild barley.
713 **Table S9.** Primer information used in this study.
714 **Table S10.** Summary of RNA-sequence experiment.
715
716 **Fig. S1.** Summary of the tonnes of certified seed production each year in UK.
717 **Fig. S2.** Additional replicates of the RNA *in situ* hybridization of *histone H4.*
718 **Fig. S3.** Frequency distributions of floret length in the lateral spikelet.
719 **Fig. S4.** Comparative basic genetic maps of *deficiens* on chromosome 2H.
720 **Fig. S5.** Alignment of deduced amino acid sequences of VRS1.
721 **Fig. S6.** Spike phenotype of *deficiens* mutants.
722 **Fig. S7.** Representation of motifs in VRS1 homologs.
723 **Fig. S8.** Scanning electron microscopy views of barley spikes.

Fig. S9. Predicted phosphorylation sites of Serine in VRS1 protein.

Fig. S10. Expression pattern of *Vrs1* and *HvHox2* in Debre Zeit 29.

Supplementary Data 1. Peptide sequences used for motif analysis.

Supplementary Data 2. GO terms of 25,122 barley genes.

Supplementary Data 3. GO terms of up-regulated genes.

Supplementary Data 4. GO terms of down-regulated genes.

731

Parsed Citations

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Arce AL, Raineri J, Capella M, Cabello JV, Chan RL (2011) Uncharacterized conserved motifs outside the HD-Zip domain in HD-Zip subfamily I transcription factors; a potential source of functional diversity. BMC Plant Biol 11: 42
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ariel FD, Manavella PA, Dezar CA, Chan RL (2007) The true story of the HD-Zip family. Trends Plant Sci 12: 419-426
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ashikari M, et al. (2005) Cytokinin oxidase regulates rice grain production. Science 309: 741-745
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res 34: W369-373
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16: 37-48
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E (2015) The Arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis 53: 474-485
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Blom N, Gammeltoft S, Brunak S (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol 294: 1351-1362
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4: 1633-1649
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bommert P, Nagasawa NS, Jackson D (2013) Quantitative variation in maize kernel row number is controlled by the FASCIATED EAR2 locus. Nat Genet
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bothmer R, Jacobsen N, Baden C, Jorgensen R, Linde-Laursen I (1995) An ecogeographical study of the genus Hordeum, Ed 2nd. IBPGR, Rome
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bull H, et al. (2017) Barley SIX-ROWED SPIKE3 encodes a putative Jumonji C-type H3K9me2/me3 demethylase that represses lateral spikelet fertility. Nat Commun
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Comadran J, et al. (2012) Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nat Genet
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127: 1309-1321
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Druka A, et al. (2011) Genetic dissection of barley morphology and development. Plant Physiol 155: 617-627
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res 38: W64-70
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Fujiwara S, Wang L, Han L, Suh SS, Salome PA, McClung CR, Somers DE (2008) Post-translational regulation of the Arabidopsis circadian clock through selective proteolysis and phosphorylation of pseudo-response regulator proteins. J Biol Chem 283: 23073-23083
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Habgood RM, Chambi JY (1984) Effects of the deficiens allele on ear development and yield in barley. Annals of Applied Biology 105: 159-166
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Han Y, Zhang C, Yang H, Jiao Y (2014) Cytokinin pathway mediates APETALA1 function in the establishment of determinate floral meristems in Arabidopsis. Proc Natl Acad Sci U S A 111: 6840-6845
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Houston K, Druka A, Bonar N, Macaulay M, Lundqvist U, Franckowiak J, Morgante M, Stein N, Waugh R (2012) Analysis of the barley bract suppression gene Trd1. Theor Appl Genet 125: 33-45
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hua L, et al. (2015) LABA1, a Domestication Gene Associated with Long, Barbed Awns in Wild Rice. Plant Cell 27: 1875-1888
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hunter T, Karin M (1992) The regulation of transcription by phosphorylation. Cell 70: 375-387
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14: R36
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kirby EJM, Appleyard M (1981) Cereal development guide, Ed 1st. Cereal Unit, Kenilworth
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Komatsuda T, et al. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc Natl Acad Sci U S A 104: 1424-1429
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Koppolu R, et al. (2013) Six-rowed spike4 (Vrs4) controls spikelet determinacy and row-type in barley. Proc Natl Acad Sci U S A 110: 13198-13203
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33: 1870-1874

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg LA (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174-181
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Li Y, et al. (2011) Natural variation in GS5 plays an important role in regulating grain size and yield in rice. Nat Genet 43: 1266-1269
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Liao Y, Smyth GK, Shi W (2013) The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res 41: e108
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Luo J, et al. (2013) An-1 encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. Plant Cell 25: 3360-3376
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mascher M, et al. (2017) A chromosome conformation capture ordered sequence of the barley genome. Nature 544: 427-433
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Poursarebani N, et al. (2015) The Genetic Basis of Composite Spike Form in Barley and 'Miracle-Wheat'. Genetics 201: 155-165
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Powers L (1936) The Nature of the Interaction of Genes Affecting Four Quantitative Characters in a Cross between Hordeum Deficiens and Hordeum Vulgare. Genetics 21: 398-420
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ramsay L, et al. (2011) INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. Nat Genet 43: 169-172
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139-140
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Saisho D, Pourkheirandish M, Kanamori H, Matsumoto T, Komatsuda T (2009) Allelic variation of row type gene Vrs1 in barley and implication of the functional divergence. Breed Sci 59: 621-628
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sakuma S, et al. (2013) Divergence of expression pattern contributed to neofunctionalization of duplicated HD-Zip I transcription factor in barley. New Phytol 197: 939-948
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sakuma S, Pourkheirandish M, Matsumoto T, Koba T, Komatsuda T (2010) Duplication of a well-conserved homeodomain-leucine zipper transcription factor gene in barley generates a copy with more specific functions. Funct Integr Genomics 10: 123-133
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sarid-Krebs L, Panigrahi KC, Fornara F, Takahashi Y, Hayama R, Jang S, Tilmes V, Valverde F, Coupland G (2015) Phosphorylation of CONSTANS and its COP1-dependent degradation during photoperiodic flowering of Arabidopsis. Plant J 84: 451-463
Pubmed: Author and Title

CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M (2008) Deletion in a gene associated with grain size increased yields during rice domestication. Nat Genet 40: 1023-1028
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Simmonds J, Scott P, Brinton J, Mestre TC, Bush M, Del Blanco A, Dubcovsky J, Uauy C (2016) A splice acceptor site mutation in TaGW2-A1 increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. Theor Appl Genet 129: 1099-1112
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet 39: 623-630
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Song XJ, et al. (2015) Rare allele of a previously unidentified histone H4 acetyltransferase enhances grain weight, yield, and plant biomass in rice. Proc Natl Acad Sci U S A 112: 76-81
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene tb1. Nat Genet 43: 1160-1163
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sugiyama N, Nakagami H, Mochida K, Daudi A, Tomita M, Shirasu K, Ishihama Y (2008) Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in Arabidopsis. Mol Syst Biol 4: 193
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

van Esse GW, Walla A, Finke A, Koornneef M, Pecinka A, von Korff M (2017) Six-Rowed Spike3 (VRS3) Is a Histone Demethylase That Controls Lateral Spikelet Development in Barley. Plant Physiol 174: 2397-2408
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang S, et al. (2012) Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet 44: 950-954
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wendt T, Holme I, Dockter C, Preuss A, Thomas W, Druka A, Waugh R, Hansson M, Braumann I (2016) HvDep1 Is a Positive Regulator of Culm Elongation and Grain Size in Barley and Impacts Yield in an Environment-Dependent Manner. PLoS One 11: e0168924
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Woodward RW (1947) The lh, l, i alleles in Hordeum deficiens genotypes of barley. J. Am. Soc. Agron. 39: 474-482
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Woodward RW (1949) The inheritance of fertility in the lateral florets of the four barley groups. Agron J. 41: 317-322
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Youssef HM, et al. (2017) VRS2 regulates hormone-mediated inflorescence patterning in barley. Nat Genet 49: 157-161
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Youssef HM, Koppolu R, Rutten T, Korzun V, Schweizer P, Schnurbusch T (2014) Genetic mapping of the labile (lab) gene: a recessive locus causing irregular spikelet fertility in labile-barley (Hordeum vulgare convar. labile). Theor Appl Genet 127: 1123-1131
Pubmed: Author and Title
CrossRef: Author and Title